

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- |                                     |  |
|-------------------------------------|--|
| n/a                                 | Confirmed  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A description of all covariates tested   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. $F$ , $t$ , $r$ ) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted<br><i>Give <math>P</math> values as exact values whenever suitable.</i>                            |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Estimates of effect sizes (e.g. Cohen's $d$ , Pearson's $r$ ), indicating how they were calculated   |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

### Software and code

Policy information about [availability of computer code](#)

Data collection	Deep targeted sequencing data: Ion Torrent Suite Software (Version 5.0.7). SNP arrays: Illumina GenomeStudio (Version 2011.1.0.24550).
Data analysis	R (version 3.6.0) was used for data analysis. R packages: minfi, ASCAT, dNdSCV, ape, InfiniumPurify, survey, ggplot2. Custom algorithms have been deposited to <a href="https://github.com/xtmgah/EAGLE_LUAD">https://github.com/xtmgah/EAGLE_LUAD</a> . Torrent variant caller (version 5.0.9).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The targeted sequencing data have been deposited in SRA through dbGaP under the accession number phs001169.v2.p1. The SNP array and methylation array data have been deposited in dbGaP under the same accession number.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Tumor samples from 84 patients were confirmed as primary lung adenocarcinoma. 2 to 11 samples from each tumor were analyzed. Sample size was determined by the availability of the patients recruited.
Data exclusions	Samples with low purity were excluded from the study. No predetermined data exclusions based on age, sex or other subjects' characteristics was applied.
Replication	No replication was applied to individual samples. But multiple regions from the same tumor were profiled.
Randomization	We did not investigate the effect of different treatments, so it is not applicable to this study.
Blinding	We did not investigate the effect of different treatments, so it is not applicable to this study.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	Age, gender, smoking status, tumor stage, survival status and time to death, metastatic status and time to metastasis
Recruitment	The patients were recruited in the EAGLE study, a population-based case-control study conducted in Italy between 2002 and 2005.
Ethics oversight	The study protocol was approved by the Institutional Review Board of the US National Cancer Institute and the involved institutions in Italy. Informed consent was obtained for all subjects prior to study participation.

Note that full information on the approval of the study protocol must also be provided in the manuscript.